# Product of group languages

S.W. Margolis and J.-E. Pin

Concatenation product is, together with Kleene's star operation, one of the most fascinating operations on recognizable (= rational, regular) languages. The study of this operation produced numerous fundamental results like Schützenberger's theorem on star-free languages, Brzozowski's results on the dot depth hierarchy, Simon's theorem on piecewiese testable languages, Straubing's characterization of varieties closed under product, results of Mc Naughton and Thomas on the connexion with first order logic, etc. It also had a considerable influence on the rest of the theory and many algebraic tools were originally introduced to produce better proofs of old or new results.

In fact, it appears that the trully fundamental operation is not exactly the usual concatenation product but a variant of it, that consists to associate to languages $L_0$, $L_1$, ..., $L_n$ the language $L_0a_1L_1a_2 \cdots a_nL_n$ where $a_1$, $a_2$, ..., $a_n$ are given letters of the alphabet. Notice that this operation is not mysterious at all. It is used for instance to obtain a rational expression associated to a finite automaton in the classical algorithm of Mc Naughton and Yamada. Therefore, in this paper, the term "product" will refer to this variant of concatenation product.

With this operation in hand, it is not difficult to construct hierarchies of recognizable languages. Start with a boolean algebra of languages: this will be the level 0 of our hierarchy. Then define level $n + 1$ as the boolean algebra generated by products (in the new sense) of languages of level $n$. If you start with the trivial boolean algebra $\{\emptyset, A^*\}$, you obtain Straubing's hierarchy. If you start with endwise testable (or "generalized definite") languages, you get Brzozowski's hierarchy, also called dot-depth hierarchy.

The aim of this paper is to study the hierarchy whose level 0 consists of all group languages. In this case, the union $\mathcal{U}$ of all levels of the hierarchy is the closure of group languages under product and boolean operations. Our first result shows that $\mathcal{U}$ is a decidable variety of languages. That is, given a recognizable language $L$, one can decide whether $L$ belongs to $\mathcal{U}$ or not. Our second result states that our hierarchy is strict. In fact, this result still holds if one takes as level 0 an arbitrary subvariety of the variety of group languages.

The rest of the paper is devoted to the study of level 1. It turns out that this variety of languages, and the corresponding variety of monoids $\diamondsuit\mathbf{G}$, appear in many different contexts. First, $\diamondsuit\mathbf{G}$ is exactly the variety $\mathbf{J} * \mathbf{G}$ generated by all semidirect products of a $\mathcal{J}$-trivial monoid by a group. This result is interesting because $\mathbf{J}$ is also the first level of Straubing's hierarchy $\mathbf{V}_n$. Thus, at least for levels 0 and 1, the operation $\mathbf{V} \to \mathbf{V} * \mathbf{G}$ is the bridge between Straubing's hierarchy and our hierarchy. Similarly, it is known [14] that the operation $\mathbf{V} \to \mathbf{V} * \mathbf{LI}$ is the bridge between Straubing's hierarchy and Brzozowski's hierarchy ($\mathbf{LI}$ denotes the first level of Brzozowski's hierarchy). Second $\diamondsuit\mathbf{G}$ is also the variety generated by powermonoids of groups. In fact we first prove the language counterpart of this result: a language has level $\leq 1$ in our hierarchy if and only if it belongs to the boolean algebra generated by all languages of the form $L\varphi$, where $L$ is a group language and $\varphi$ is a length preserving morphism (strictly alphabetic, letter to letter morphism). Finally, languages of level 1 arise in the study of the finite group topology for the free monoid [7].

An important problem is to know whether the first level of our hierarchy is a *decidable* variety. The answer is positive in the case of Straubing's hierarchy and Brzozowski's hierarchy, and thus we hope a positive answer also in our case. The discussion of this problem motivated the introduction of a new variety of monoids, denoted by **BG**. **BG** is the variety of all monoids $M$ such that, for every idempotent $e, f \in M$, $efe = e$ implies $e = f$. Several equivalent definitions are given in the paper. In particular, we show that a monoid $M$ is in **BG** if and only if the submonoid generated by all idempotents of $M$ belongs to **J**. We also prove that **BG** is generated by all monoids that are, in some sense, extensions of a group by a monoid of **J**. Finally, $\Diamond\mathbf{G}$ is contained in **BG** but we don't know if this inclusion is strict or not. If $\Diamond\mathbf{G} = \mathbf{BG}$, then $\Diamond\mathbf{G}$ is decidable. If $\Diamond\mathbf{G} \neq \mathbf{BG}$ (our conjecture) then $\Diamond\mathbf{G}$ should satisfy some mysterious equations. Let us give some details: let $G$ be a group and let $\mathcal{P}(G)$ be the powergroup of $G$. Then the idempotents of $\mathcal{P}(G)$ are exactly the subgroups of $G$, and if $H$ and $K$ are subgroups of $G$, $HKH = H$ implies $H = K$: this proves that $\mathcal{P}(G)$ belongs to **BG**. However, it is hard to imagine some others algebraic conditions that are true in $\mathcal{P}(G)$ for *any finite group $G$*.

# 1 Preliminaries.

In this paper, all semigroups are finite except in the case of free monoids. A language is a *group language* if and only if its syntactic monoid is a finite group.

Given a monoid $M$, we denote by $E(M)$ the set of idempotents of $M$ and by $EG(M)$ the subsemigroup of $M$ generated by $E(M)$. For every $x \in M$, $x^\omega$ denotes the (unique) idempotent of the subsemigroup of $M$ generated by $x$. Notice that if $n_x$ denotes the smallest positive such that $x^{n_x}$ is idempotent, then one can take $\omega = \mathrm{lcm}\{n_x \mid x \in M\}$ and thus $\omega$ does not depend on the choice of $x$.

$\mathcal{P}(M)$ denotes the *powermonoid* of $M$, that is the monoid of subsets of $M$ under the multiplication given by $XY = \{xy \mid x \in X \text{ and } y \in Y\}$.

A *variety of monoids* is a class of monoids closed under taking submonoids, quotients and finite direct products. We refer to [1, 3, 6] for more details on varieties. A *variety of groups* is a variety of monoids whose elements are groups. Here is the list of the varieties used in this paper

- **A** aperiodic (or group-free) monoids
- **G** groups
- $\mathbf{G}_p$ $p$-groups
- **I** trivial variety $\{1\}$
- **J** $\mathcal{J}$-trivial monoids (monoids $M$ such that $MaM = MbM$ implies $a = b$)
- **R** $\mathcal{R}$-trivial monoids (monoids $M$ such that $aM = bM$ implies $a = b$)

To each variety of monoids $\mathbf{V}$, one associates the corresponding variety of languages $\mathcal{V}$. For each alphabet $A$, $A^*\mathcal{V}$ is the set of all (recognizable) languages $L$ of $A^*$ whose syntactic monoid $M(L)$ belongs to $\mathbf{V}$. The variety theorem [1] states that the correspondence $\mathbf{V} \to \mathcal{V}$ is one to one.

There are a number of operations defined on varieties. Given a variety of monoids $\mathbf{V}$, $\mathbf{EV}$ denotes the variety of all monoids $M$ such that $GE(M) \in \mathbf{V}$. $\mathbf{PV}$ denotes the variety of monoids generated by all monoids of the form $\mathcal{P}(M)$ where $M \in \mathbf{V}$. Similarly, $\Diamond\mathbf{V}$ denotes the variety generated by all Schützenberger products [5, 13] $\Diamond_k(M_1, \ldots, M_k)$ where $k > 0$ and $M_1, \ldots, M_k \in \mathbf{V}$. The following result was proved in [5].

**Proposition 1.1** *Let $\mathcal{V}$ (resp. $\Diamond\mathcal{V}$) be the variety of languages corresponding to $\mathbf{V}$ (resp. $\Diamond\mathbf{V}$). Then for every alphabet $A$, $A^*\Diamond\mathcal{V}$ is the boolean algebra generated by all languages of the form $L_0a_1L_1 \cdots a_kL_k$ where $k \geq 0$, $a_1, \ldots, a_k \in A$, $L_0, \ldots, L_k \in A^*\mathcal{V}$.*

More generally we define $\Diamond^n\mathbf{V}$ by induction as follows: $\Diamond^0\mathbf{V} = \mathbf{V}$ and $\Diamond^{n+1}\mathbf{V} = \Diamond(\Diamond^n\mathbf{V})$. In particular, the hierarchy $\Diamond^n(\mathbf{I}) = \mathbf{V}_n$ is known as Straubing's hierarchy. The next proposition summarizes some known results on $\mathbf{V}_n$.

**Proposition 1.2**

(a) $\mathbf{V}_0 = \mathbf{I}$, $\mathbf{V}_1 = \mathbf{J}$ *and* $\mathbf{V}_2 = \mathbf{PJ}$.

(b) $\bigcup_{n\geq 0} \mathbf{V}_n = \mathbf{A}$ *and* $\mathbf{V}_n \subsetneq \mathbf{V}_{n+1}$ *for every* $n > 0$.

Given a variety of groups $\mathbf{H}$, $\overline{\mathbf{H}}$ denotes the variety of all monoids whose groups are elements of $\mathbf{H}$.

Let $\mathbf{V}$ and $\mathbf{W}$ be two varieties of monoids. Then $\mathbf{V} * \mathbf{W}$ denotes the variety of monoids generated by all semidirect products $M * N$ where $M \in \mathbf{V}$ and $N \in \mathbf{W}$.

Finally $\mathbf{V}^{-1}\mathbf{W}$ denotes the variety generated by all monoids $M$ such that there exists a $\mathbf{V}$-morphism $\varphi : M \to N$ where $N \in \mathbf{W}$. Recall that $\varphi$ is a $\mathbf{V}$-morphism whenever for every subsemigroup $S$ of $N$, $S \in \mathbf{V}$ implies $S\varphi^{-1} \in \mathbf{V}$.

# 2 Concatenation of group languages

In this section we introduce hierarchies analoguous to Straubing's and Brzozowski's hierarchy and we show that these hierarchies are infinite.

Let $\mathbf{H}$ be a variety of groups and let $\mathcal{H}$ be the corresponding variety of languages. We construct a hierarchy of varieties $\Diamond^n\mathcal{H}$ ($n \geq 0$) as follows: $\Diamond^0\mathcal{H} = \mathcal{H}$ and for every alphabet $A$, $A^*\Diamond^{n+1}\mathcal{H}$ is the boolean algebra generated by all languages of the form $L_0a_1L_1 \cdots a_kL_k$ where $L_0, \ldots, L_k \in A^*\Diamond^n\mathcal{H}$ and $a_1, \ldots, a_k \in A$. Finally set $\Diamond^\infty\mathcal{H} = \bigcup_{n\geq 0} \Diamond^n\mathcal{H}$. Notice that $\Diamond^\infty\mathcal{H}$ is the smallest variety closed under product containing $\mathcal{H}$. Then we have:

**Theorem 2.1** *For every $n \geq 0$, $\Diamond^n\mathcal{H}$ corresponds to the variety of monoids $\Diamond^n\mathbf{H}$. Furthermore $\Diamond^\infty\mathcal{H}$ corresponds to $\mathbf{A}^{-1}\mathbf{H} = \mathbf{A} * \mathbf{H}$.*

The first part of the statement was proved in [5]. Furthermore $\Diamond^\infty\mathcal{H}$ is the closure of $\mathcal{H}$ under product. Thus by a theorem of Straubing [11], it corresponds to $\mathbf{A}^{-1}\mathbf{H}$. Finally the equality $\mathbf{A}^{-1}\mathbf{H} = \mathbf{A} * \mathbf{H}$ was also proved in [11].

**Corollary 2.2** *One can decide whether a given recognizable language belongs to the variety $\diamondsuit^\infty \mathcal{G}$.*

**Proof.** By Theorem 2.1, we have $L \in A^*\diamondsuit^\infty\mathcal{G}$ if and only if $M(L) \in \mathbf{A}^{-1}\mathbf{G}$. But a result of Karnofsky and Rhodes [2] shows that this last condition is decidable.

Our next result shows that our hierarchies are always infinite.

**Theorem 2.3** *For every variety of groups* $\mathbf{H}$*, the hierarchy* $\diamondsuit^n\mathbf{H}$ *is strict. That is, for every* $n \geq 0$, $\diamondsuit^n\mathbf{H} \subsetneq \diamondsuit^{n+1}\mathbf{H}$.

The proof uses a result of independent interest.

**Proposition 2.4** *Let* $\mathbf{V}$ *be a variety of monoids. Then* $\mathbf{V}^{-1}\mathbf{G} \subset \mathbf{EV}$.

**Proof.** Let $G \in \mathbf{H}$ be a group and let $\pi : M \to G$ be a $\mathbf{V}$-morphism. By definition, for every subsemigroup $S$ of $G$, $S \in \mathbf{V}$ implies $S\pi^{-1} \in \mathbf{V}$. In particular, $1\pi^{-1} \in \mathbf{V}$. Now let $e \in E(M)$. Then $e\pi$ is also an idempotent in $G$ and thus $e\pi = 1$. It follows $E(M) \subset 1\pi^{-1}$ and since $1\pi^{-1}$ is a semigroup, $GE(M) \subset 1\pi^{-1}$. Therefore $GE(M) \in \mathbf{V}$ and hence $M \in \mathbf{EV}$. It follows that $\mathbf{V}^{-1}\mathbf{G} \subset \mathbf{EV}$. □

Following Straubing [13] we define for each $n > 0$ a variety of monoids $\mathbf{W}_n$ by the equations $x^\omega = x^{\omega+1}$ and $(xy)^\omega\alpha^{n-1} = (yx)^\omega\alpha^{n-1}$ where $\alpha : \{x, y\}^* \to \{x, y\}^*$ is the morphism defined by $x\alpha = (xy)^\omega x(xy)^\omega$ and $y\alpha = (xy)^\omega y(xy)^\omega$. The next proposition summarizes some results of [13].

**Proposition 2.5**

(a) $\bigcup_{n>0} \mathbf{W}_n = \mathbf{A}$

(b) *For each* $n > 0$, $\mathbf{W}_n \subset \mathbf{W}_{n+1}$

(c) *For each variety of groups* $\mathbf{H}$, $\diamondsuit^n\mathbf{H} \subset \mathbf{W}_n^{-1}\mathbf{H}$.

We are now ready to prove Theorem 2.3. Assume that $\diamondsuit^n\mathbf{H} = \diamondsuit^{n+1}\mathbf{H}$ for some $n \geq 0$. Then by induction on $p$, $\diamondsuit^n\mathbf{H} = \diamondsuit^{n+p}\mathbf{H}$ for every $p > 0$ and thus, by Theorem 2.1, we have $\diamondsuit^n\mathbf{H} = \bigcup_{p\geq 0} \diamondsuit^{n+p}\mathbf{H} = \mathbf{A}^{-1}\mathbf{H}$. Now, by Proposition 2.4 and 2.5 c) we obtain

$$\mathbf{A} \subset \mathbf{A}^{-1}\mathbf{H} \subset \diamondsuit^n\mathbf{H} \subset \mathbf{W}_n^{-1}\mathbf{H} \subset \mathbf{EW}_n$$

Now by Proposition 2.5 a) and b) there exists a monoid $M \in \mathbf{A} \setminus \mathbf{W}_n$. Furthermore one can show (proof omitted) that every aperiodic monoid $M$ divides an idempotent generated aperiodic monoid $N$. Thus if $M \in \mathbf{A} \setminus \mathbf{W}_n$, $N \in \mathbf{A} \setminus \mathbf{W}_n$ and since $N = GE(N)$ we also have $N \in \mathbf{A} \setminus \mathbf{EW}_n$, a contradiction. □

The last result of this section relates our hierarchies to the hierarchy of Straubing $\mathbf{V}_n$.

**Theorem 2.6** *For every* $n \geq 0$ *and for every variety of groups* $\mathbf{H}$, $\mathbf{V}_n * \mathbf{H} \subset \diamondsuit^n\mathbf{H}$.

As we shall see in the next section, this inclusion is an equality if $\mathbf{H} = \mathbf{G}$ or $\mathbf{G}_p$, and if $n = 0$ or 1. However we don't know if this equality still holds in the general case. Theorem 2.6 will be obtained as a consequence of a more general result, the proof of which is a good illustration of the connexions between languages and semigroups.

**Theorem 2.7** *For any variety of monoids* $\mathbf{V}$ *and* $\mathbf{W}$, $(\Diamond\mathbf{V}) * \mathbf{W} \subset \Diamond(\mathbf{V} * \mathbf{W})$.

**Proof.** Let $\mathcal{V}$, $\mathcal{W}$, $\mathcal{U}$, $\mathcal{X}$, $\mathcal{S}$ and $\mathcal{T}$ be the varieties of languages corresponding to $\mathbf{V}$, $\mathbf{W}$, $\mathbf{V} * \mathbf{W}$, $\Diamond\mathbf{V}$, $(\Diamond\mathbf{V}) * \mathbf{W}$ and $\Diamond(\mathbf{V} * \mathbf{W})$ respectively. By Eilenberg theorem, it suffices to show that $\mathcal{S} \subset \mathcal{T}$. Let $L \in A^*\mathcal{S}$. Then $L$ is recognized by a wreath product $M \circ N$ where $M \in \Diamond\mathbf{V}$ and $N \in \mathbf{W}$. That is, there exists a morphism $\eta : A^* \to M \circ N$ such that $L = L\eta\eta^{-1}$. Let $\pi : M \circ N \to N$ be the natural projection and let $\varphi = \eta\pi : A^* \to N$. Finally set $B = N \times A$ and let $\sigma : A^* \to B^*$ be the sequential function defined by

$$(a_1a_2 \cdots a_n)\sigma = (1, a_1)(a_1\varphi, a_2) \cdots ((a_1 \cdots a_{n-1})\varphi, a_n)$$

Then, by the *wreath product principle* [10], $L$ is a finite union of languages of the form $X \cap Y\sigma^{-1}$ where $X \in A^*\mathcal{W}$ and $Y \in B^*\mathcal{X}$. Since $\mathbf{W} \subset \Diamond(\mathbf{V} * \mathbf{W})$ we have $X \in A^*\mathcal{T}$. Furthermore, $A^*\mathcal{T}$ is a boolean algebra by definition and thus it suffices to show that $Y\sigma^{-1} \in A^*\mathcal{T}$. Now since $Y \in B^*\mathcal{X}$, $Y$ is a boolean combination of languages of the form $K = L_0b_1L_1b_2 \cdots b_kL_k$ where $L_0, \ldots, L_k \in B^*\mathcal{V}$ and $b_1, \ldots, b_k \in B$. Since $\sigma^{-1}$ commutes with boolean operations, it suffices to show that such a language $K$ is in $A^*\mathcal{T}$. Set $b_i = (n_i, a_i)$ for $1 \leq i \leq k$. Then we have

$$\begin{aligned} K &= \{c_1 \cdots c_p \in A^* \mid (c_1 \cdots c_p)\sigma \in L_0b_1L_1 \cdots b_kL_k\} \\ &= \{c_1 \cdots c_p \in A^* \mid (1, c_1)(c_1\varphi, c_2) \cdots ((c_1 \cdots c_{p-1})\varphi, c_p) \in L_0b_1L_1 \cdots b_kL_k\} \end{aligned}$$

For $1 \leq r \leq k$, let $\sigma_r : A^* \to B^*$ be the (sequential) function defined by:

$$(a_1a_2 \cdots a_n)\sigma_r = (h_r, a_1)(h_r(a_1\varphi), a_2) \cdots (h_r(a_1a_2 \cdots a_{n-1}\varphi, a_n)$$

where $h_r = n_r(a_r\varphi)$, and let $\sigma_0 = \sigma$. Now a simple calculation shows that

$$\begin{aligned} K = \{c_1 \cdots c_p \in A^* \mid\ & \exists i_1 < i_2 \cdots < i_k \text{ such that} \\ & (1) \quad ((c_1 \cdots c_{i_r - 1})\varphi, c_{i_r}) = b_r = (n_r, a_r) \text{ for } 1 \leq r \leq k \text{ and} \\ & (2) \quad \begin{cases} (c_1 \cdots c_{i_r - 1})\sigma_0 \in L_0 \\ (c_{i_r + 1} \cdots c_{i_{r+1} - 1})\sigma_r \in L_r \quad \text{for } 1 \leq r \leq k-1 \quad \} \\ (c_{i_k + 1} \cdots c_p)\sigma_k \in L_k \end{cases} \end{aligned}$$

It follows

$$\begin{aligned} K = \{u \in A^* \mid\ & \text{there exists a factorization } u = u_0b_1u_1 \cdots b_ku_k \text{ such that} \\ & (1) \quad (u_0a_1u_1 \cdots a_ru_r)\varphi = n_{r+1} \text{ for } 0 \leq r \leq k-1 \\ & (2) \quad u_r\sigma_r \in L_r \text{ for } 0 \leq r \leq k \ \} \\ = \{u \in A^* \mid\ & \text{there exists a factorization } u = u_0b_1u_1 \cdots b_ku_k \text{ such that} \\ & (1) \quad h_r(u_r\varphi) = n_{r+1} \text{ for } 0 \leq r \leq k-1 \text{ (where } h_0 = 1 \text{ and} \\ & \qquad\qquad h_r = n_r(a_r\varphi) \text{ for } 0 \leq r \leq k-1) \\ & (2) \quad u_r\sigma_r \in L_r \\ = S_0b_1S_1 & \cdots b_kS_k \end{aligned}$$

where $S_r = L_r\sigma_r^{-1} \cap (h_r^{-1}n_{r+1})\varphi^{-1}$ for $0 \leq r \leq k-1$ and $S_k = L_k\sigma_k^{-1}$. Now, $(h_r^{-1}n_{r+1})\varphi^{-1}$ is recognized by $N$ and thus belongs to $A^*\mathcal{W}$, and a fortiori to $A^*\mathcal{U}$ since $\mathbf{W} \subset \mathbf{V}*\mathbf{W}$. Furthermore, $\sigma_r$ is a sequential function whose syntactic monoid belongs to $\mathbf{W}$. Since $L_r \in A^*\mathcal{V}$, we have $M(L_r\sigma_r^{-1}) \in \mathbf{V}*\mathbf{W}$ and thus $L_r\sigma_r^{-1} \in A^*\mathcal{U}$. Finally $S_r \in A^*\mathcal{U}$ and thus $K \in A^*\mathcal{T}$ by construction. □

Notice that the inclusion $(\Diamond\mathbf{V}) * \mathbf{W} \subset \Diamond(\mathbf{V} * \mathbf{W})$ may be proper. For instance, if $\mathbf{V} = \mathbf{I}$, $\mathbf{W} = \mathbf{J}$, we have $(\Diamond\mathbf{V}) * \mathbf{W} = \mathbf{J} * \mathbf{J} = \mathbf{R}$ and $\Diamond(\mathbf{V} * \mathbf{W}) = \Diamond(\mathbf{J}) = \mathbf{PJ} = \mathbf{PR}$.

**Corollary 2.8** *For every $n \geq 0$ and for every variety of monoids $\mathbf{W}$, $\mathbf{V}_{n+1} * \mathbf{W} \subset \Diamond(\mathbf{V}_n * \mathbf{W})$*

**Proof.** This is a consequence of Theorem 2.7 since $\mathbf{V}_{n+1} = \Diamond(\mathbf{V}_n)$. □

**Corollary 2.9** *For every $n \geq 0$ and for every variety of monoids $\mathbf{W}$, $\mathbf{V}_n * \mathbf{W} \subset \Diamond^n(\mathbf{W})$.*

**Proof.** By induction on $n$. □

# 3 The level 1

In this section we investigate the properties of the first level of our hierarchies. It turns out that this first level is related to various problems of language theory and that the corresponding variety of monoids admits several different characterizations.

Let us first recall the precise definition. Let $\mathbf{H}$ be a variety of groups and let $\mathcal{H}$ be the corresponding variety of languages. In the sequel, we shall mainly consider the case $\mathbf{H} = \mathbf{G}$ or $\mathbf{G}_p$. Then $\Diamond\mathcal{H}$ is the variety of languages defined as follows. For every alphabet $A$, $A^*\Diamond\mathcal{H}$ is the boolean algebra generated by all languages of the form $L_0a_1 \cdots a_kL_k$ where $k \geq 0$, $L_0, \ldots, L_k \in A^*\mathcal{H}$ and $a_1, \ldots, a_k \in A$. Our first result gives an alternative description in the case $\mathbf{H} = \mathbf{G}$ or $\mathbf{G}_p$.

**Theorem 3.1** *Let $\mathbf{H} = \mathbf{G}$ or $\mathbf{G}_p$ for some prime number $p$. Then for every alphabet $A$, $A^*\Diamond\mathcal{H}$ is the boolean algebra generated by all languages of the form $L\varphi$ where $L \in B^*\mathcal{H}$ and $\varphi : B^* \to A^*$ is a length preserving morphism.*

It is known [8, 12] that if $\mathbf{V}$ is the variety of monoids corresponding to a variety of languages $\mathcal{V}$, then $\mathbf{PV}$ corresponds to the variety $\mathcal{W}$ defined as follows: for every alphabet $A$, $A^*\mathcal{W}$ is the boolean algebra generated by all languages of the form $K\varphi$, where $K \in B^*\mathcal{V}$ and $\varphi : B^* \to A^*$ is a length preserving morphism. Therefore, Theorem 3.1 is equivalent to state that $\mathbf{PH} = \Diamond\mathbf{H}$. In fact we prove a slightly more complete result.

**Theorem 3.2** *Let $\mathbf{H} = \mathbf{G}$ or $\mathbf{G}_p$ for some prime number $p$. Then $\Diamond\mathbf{H} = \mathbf{PH} = \mathbf{J} * \mathbf{H}$.*

**Proof.** We treat the case $\mathbf{H} = \mathbf{G}_p$ only but the case $\mathbf{H} = \mathbf{G}$ is analoguous. Let $\mathcal{W}$ be the variety corresponding to $\mathbf{PH}$. We first show that $\Diamond\mathcal{G}_p \subset \mathcal{W}$ (and thus $\Diamond\mathbf{H} \subset \mathbf{PH}$ by the variety theorem). It suffices to prove that every language of the form $L = L_0 a_1 L_1 \cdots a_k L_k$ — where $L_0, \ldots, L_k \in A^*\mathcal{G}_p$ and $a_1, \ldots, a_k \in A$ — can be written as $K\pi$ — where $K \in B^*\mathcal{G}_p$ and $\pi : B^* \to A^*$ is length preserving.

Let $\bar{A}$ be a copy of $A$ and let $\pi : (A \cup \bar{A})^* \to A^*$ be the length preserving morphism defined by $a\pi = \bar{a}\pi = a$ for every $a \in A$. Set $K_i = L_i\pi^{-1}$ for $0 \leq i \leq k$ and let $K$ be the set of all words $u \in (A \cup \bar{A})^*$ whose number of factorizations of the form $u_0\bar{a}_1u_1 \cdots \bar{a}_k u_k$ — with $u_0 \in K_0, \ldots, u_k \in K_k$ — is congruent to 1 modulo $p$. One can show that the syntactic monoid of $K$ is a $p$-group and hence $K \in (A\cup\bar{A})^*\mathcal{G}_p$ . We claim that $K\pi = L$. Indeed we have $K \subset K_0\bar{a}_1K_1 \cdots \bar{a}_kK_k$ and thus $K\pi \subset (K_0\bar{a}_1K_1 \cdots \bar{a}_kK_k)\pi = L_0\bar{a}_1L_1 \cdots \bar{a}_kL_k = L$. Conversely let $u \in L$. Then $u$ admits a factorization of the form $u = u_0a_1u_1 \cdots a_ku_k$ where $u_0 \in L_0, \ldots, u_k \in L_k$. Each $u_i$ can be viewed as a word of $(A \cup \bar{A})^*$ that we shall denote by $v_i$ to avoid any confusion. Thus $v_i\pi = u_i$ by definition and hence $v_i \in K_i$ for $0 \leq i \leq k$. Therefore $v = v_0\bar{a}_1v_1 \cdots \bar{a}_kv_k \in K_0\bar{a}_1K_1 \cdots \bar{a}_kK_k$ and $v\pi = u$. Furthermore $v_0\bar{a}_1v_1 \cdots \bar{a}_kv_k$ is the unique factorization of $v$ in $K_0\bar{a}_1K_1 \cdots \bar{a}_kK_k$ and hence $v \in K$. Thus $L \subset K\pi$ and finally $L = K\pi$ as required. □

The next step of the proof is the following proposition

**Proposition 3.3** *For every variety of groups* $\mathbf{H}$, $\mathbf{PH} \subset \mathbf{J} * \mathbf{H}$.

Recall that $U_1$ denotes the monoid $\{0, 1\}$ under the usual multiplication of integers. Let $G$ be a group and let $\mathcal{P}(G)$ (resp. $\mathcal{P}'(G)$, $\mathcal{P}_1(G)$) be the monoid of all subsets (resp. non empty subsets, subsets containing the identity of $G$) under the usual multiplication of subsets. It suffices to show that if $G \in \mathbf{H}$, then $\mathcal{P}(G) \in \mathbf{J} * \mathbf{H}$. First $\mathcal{P}(G)$ is a quotient of $\mathcal{P}'(G) \times U_1$. Indeed, let $\varphi : \mathcal{P}'(G) \times U_1 \to \mathcal{P}(G)$ be the function defined by $(X, 0)\varphi = \emptyset$ and $(X, 1)\varphi = X$ for every $X \in \mathcal{P}'(G)$. Then $\varphi$ is a surjective morphism. Now we have $U_1 \in \mathbf{J} \subset \mathbf{J} * \mathbf{H}$ and since a variety is closed under direct product, it suffices to show that $\mathcal{P}'(G)\mathbf{J} * \mathbf{H}$. Next we claim that $\mathcal{P}'(G)$ is a quotient of a semidirect product $\mathcal{P}_1(G) * G$. Define an action $G \times \mathcal{P}_1(G) \to \mathcal{P}_1(G)$ by setting

$$g{\cdot}X = gXg^{-1} \text{ for every } g \in G \text{ and } X \in \mathcal{P}_1(G)$$

This defines a semidirect product $\mathcal{P}_1(G) * G$. Furthermore, the function $\psi : \mathcal{P}_1(G) * G \to \mathcal{P}'(G)$ defined by $(X, g)\psi = Xg$ is a surjective morphism and this proves the claim.

Finally, we show that $\mathcal{P}_1(G) \in \mathbf{J}$. Let $X, Y \in \mathcal{P}_1(G)$ be two $\mathcal{J}$-related elements. Then $AXB = Y$ and $CYD = X$ for some $A, B, C, D \in \mathcal{P}_1(G)$. Since 1 belongs to $A$, $B$, $C$ and $D$, it follows $X = \{1\}X\{1\} \subset AXB = Y$ and $Y = \{1\}Y\{1\} \subset CYD = X$ and thus $X = Y$.

Thus $\mathcal{P}_1(G) * G \in \mathbf{J} * \mathbf{H}$ and hence $\mathcal{P}'(G) \in \mathbf{J} * \mathbf{H}$ as required. □

Notice that the inclusion of Proposition 3.3 may be strict. For instance, if $\mathbf{H}$ is the trivial variety, then $\mathbf{PH} = \mathbf{J}_1$ is strictly contained in $\mathbf{J} * \mathbf{H} = \mathbf{J}$.

We can now conclude the proof of Theorem 3.2. We have shown up to now the inclusions $\Diamond\mathbf{H} \subset \mathbf{PH} \subset \mathbf{J} * \mathbf{H}$. Furthermore, Theorem 2.6 (for $n = 1$) gives

$\mathbf{V}_1 * \mathbf{H} \subset \Diamond\mathbf{H}$, that is $\mathbf{J} * \mathbf{H} \subset \Diamond\mathbf{H}$ since $\mathbf{V}_1 = \mathbf{J}$. Therefore $\Diamond\mathbf{H} = \mathbf{PH} = \mathbf{J} * \mathbf{H}$. □

# 4 An approximation of the first level

The results of the previous section give various descriptions of the variety $\Diamond\mathbf{G}$, but none of these characterizations is effective. That is, we still don't know if there is an algorithm to decide whether a given monoid belongs to $\Diamond\mathbf{G}$ or not. During the analysis of this problem, we were led to introduce a new variety, denoted by $\mathbf{BG}$. This variety is decidable and just like $\Diamond\mathbf{G}$, it admits various interesting and natural characterizations. $\Diamond\mathbf{G}$ is contained in $\mathbf{BG}$ but we still don't know if, according to our believe, this inclusion is strict. Thus $\mathbf{BG}$ appears as a rather good "approximation" of $\Diamond\mathbf{G}$.

Let us first introduce a convenient definition. A monoid is *block-group* if and only if every regular $\mathcal{D}$-class of $M$ is a Brandt semigroup. The reader that is not familiar with semigroup theory can replace this definition by condition (3) of the next proposition, the proof of which is omitted.

**Proposition 4.1** *Let $M$ be a monoid. The following conditions are equivalent:*
(1) *$M$ is block-group.*
(2) *For every $e, f \in E(M)$, $e \mathcal{R} f$ or $e \mathcal{L} f$ implies $e = f$.*
(3) *For every $s \in M$ and $e \in E(M)$, $ese = e$ implies $es = e = se$.*
(4) *For every $e, f \in E(M)$, $efe = e$ implies $ef = e = fe$.*

Block-group monoids from a variety, denoted by $\mathbf{BG}$. Similarly, block-group monoids whose groups are elements of a given variety of groups $\mathbf{H}$ form a variety, denoted by $\mathbf{BH}$. Clearly $\mathbf{BH} = \mathbf{BG} \cap \overline{\mathbf{H}}$ and furthermore

**Theorem 4.2** *For every variety of groups* $\mathbf{H}$, $\mathbf{PH} \subset \mathbf{J} * \mathbf{H} \subseteq \Diamond\mathbf{H} \subset \mathbf{J}^{-1}\mathbf{H} \subset \mathbf{BH}$.

**Proof.** The first and second inclusions were already proved above. The third one is a consequence of Proposition 2.5 (c), with $n = 1$, since $\mathbf{W}_1 = \mathbf{J}$. Let us prove the last inclusion. First we have $\mathbf{J}^{-1}\mathbf{H} \subset \mathbf{A}^{-1}\mathbf{H} \subset \overline{\mathbf{H}}$ and $\mathbf{J}^{-1}\mathbf{H} \subset \mathbf{J}^{-1}\mathbf{G}$. since $\mathbf{BH} = \overline{\mathbf{H}} \cap \mathbf{BG}$, it suffices to show that $\mathbf{J}^{-1}\mathbf{G} \subset \mathbf{BG}$. Let $G$ be a group and let $\pi : M \to G$ be a $\mathbf{J}$-morphism. Let $e, f \in E(M)$ be such that $e \mathcal{R} f$ or $e \mathcal{L} f$. Then $e\pi = f\pi = 1$ and hence $e, f \in 1\pi^{-1}$. Since $1\pi^{-1}$ is $\mathcal{J}$-trivial, it follows $e = f$. Thus $M \in \mathbf{BG}$ by Proposition 4.1. □

The next proposition gives the equations of the variety $\mathbf{BG}$

**Proposition 4.3** *The variety* $\mathbf{BG}$ *is defined by one of the following equations*
(1) $(x^\omega y)^\omega = (yx^\omega)^\omega$,
(2) $(x^\omega y^\omega)^\omega = (y^\omega x^\omega)^\omega$
(3) $(x^\omega y^\omega)^\omega x^\omega = (x^\omega y^\omega)^\omega = y^\omega (x^\omega y^\omega)^\omega$

**Proof.** If $M \in \mathbf{BG}$, then $M$ satisfies (1). Indeed let $e = (x^\omega y)^\omega$ and $f = x^\omega$. Then $fe = e$ and hence $efe = e$. It follows $e = fe = ef$ by Proposition 4.1, that is $(x^\omega y)^\omega = (x^\omega y)^\omega x^\omega$. Similarly $(yx^\omega)^\omega = x^\omega (yx^\omega)^\omega$. But $(x^\omega y)^\omega x^\omega = x^\omega (yx^\omega)^\omega$ and thus $(x^\omega y)^\omega = (yx^\omega)^\omega$.

Equation (2) is deduced from (1) by replacing $y$ by $y^\omega$. Thus if $M$ satisfies (1), it also satisfies (2).

Similarly if $M$ satisfies (2) we have

$$
\begin{aligned}
(x^\omega y^\omega)^\omega x^\omega = (y^\omega x^\omega)^\omega x^\omega = (y^\omega x^\omega)^\omega = (x^\omega y^\omega)^\omega \\
= (x^\omega y^\omega)^\omega y^\omega = (y^\omega x^\omega)^\omega y^\omega = y^\omega (x^\omega y^\omega)^\omega
\end{aligned}
$$

and thus $M$ satisfies (3).

Finally, assume that $M$ satisfies (3) and let $e, f \in E(M)$. Then by (3) $(e^\omega f^\omega)^\omega e^\omega = (e^\omega f^\omega)^\omega = f^\omega (e^\omega f^\omega)^\omega$, that is $(ef)^\omega e = (ef)^\omega = f(ef)^\omega$. Now if $e \mathrel{\mathcal{R}} f$, we have $ef = f$ and $fe = e$ and thus $(ef)^\omega e = f^\omega e = fe = e$, $(ef)^\omega = f^\omega = f$ and finally $e = f$. Therefore $M \in \mathbf{BG}$ by Proposition 4.1 (2). □

The definition of a block-group monoid is given by a "local" condition on each regular $\mathcal{D}$-class. Our next result gives a "global" characterization.

**Proposition 4.4** *A monoid $M$ is a block-group if and only if $EG(M) \in \mathbf{J}$.*

**Proof.** Assume that $EG(M)$ is $\mathcal{J}$-trivial. Then for every $x, y \in M$, we have $(x^\omega y^\omega)^\omega x^\omega \mathrel{\mathcal{R}} (x^\omega y^\omega)^\omega$ in $EG(M)$ and thus $(x^\omega y^\omega)^\omega x^\omega = (x^\omega y^\omega)^\omega$. Similarly $(x^\omega y^\omega)^\omega = y^\omega (x^\omega y^\omega)^\omega$. Thus $M$ is block-group by Proposition 4.3.

Conversely assume that $M$ is block-group and let $e_1, \ldots, e_n \in E(M)$. Then one can prove by induction on $i$ that for every $0 \le i \le n$, $(e_1 \cdots e_n)^\omega = (e_1 \cdots e_n)^\omega e_1 \cdots e_i$ (proof omitted). Let now $x$ and $y$ be two $\mathcal{R}$-equivalent elements of $EG(M)$. Then $xe_1 \cdots e_i = y$ and $ye_{i+1} \cdots e_n = x$ for some $e_1, \ldots, e_n \in E(M)$. It follows

$$x = x(e_1 \cdots e_n) = x(e_1 \cdots e_n)^\omega = x(e_1 \cdots e_n)^\omega e_1 \cdots e_i = xe_1 \cdots e_i = y$$

Thus $EG(M)$ is $\mathcal{R}$-trivial and a dual proof shows that $EG(M)$ is $\mathcal{L}$-trivial. Therefore $EG(M)$ is $\mathcal{J}$-trivial. □

Our next result relates block-group monoids and extensions of groups by $\mathcal{J}$-trivial monoids. Although a purely algebraic proof of this result is possible, we present here a proof using context-free grammars.

**Theorem 4.5** *The following equality holds:* $\mathbf{BG} = \mathbf{J}^{-1}\mathbf{G}$.

**Proof.** Let $G$ be a group and let $\pi : M \to G$ be a $\mathcal{J}$-morphism. Let $e, f \in E(M)$ and assume that $efe = e$. Then $e\pi = f\pi = 1$. It follows that $fe \mathrel{\mathcal{L}} e \mathrel{\mathcal{R}} ef$ in $1\pi^{-1}$. But $1\pi^{-1}$ is $\mathcal{J}$-trivial and hence $fe = e = ef$. Thus $M \in \mathbf{BG}$ by Proposition 4.1. This proves the inclusion $\mathbf{J}^{-1}\mathbf{G} \subset \mathbf{BG}$.

Conversely, let $M \in \mathbf{BG}$ and let $D(M)$ be the submonoid of $M$ generated by the grammar

$$\xi \to s\xi\bar{s} + \bar{s}\xi s + \xi\xi + 1 \text{ for every } s, \bar{s} \in M \text{ such that } s\bar{s}s = s.$$

It follows from a result of Rhodes and Tilson [9] that $M \in \mathbf{J}^{-1}\mathbf{G}$ if and only if $D(M) \in \mathbf{J}$. Furthermore, Proposition 4.4 shows that $M \in \mathbf{BG}$ if and only if $EG(M) \in \mathbf{J}$. Thus it suffices to prove the following lemma

**Lemma 4.6** *$D(M) \in \mathbf{J}$ if and only if $EG(M) \in \mathbf{J}$.*

**Proof.** Let $e \in E(M)$. Then $ese = e$ and thus $\xi \to ese$ is a rule of our grammar. It follows that $\xi \xrightarrow{*} e$ and thus $e \in D(M)$. Therefore $EG(M)$ is a submonoid of $D(M)$ and hence if $D(M) \in \mathbf{J}$, then $EG(M) \in \mathbf{J}$.

Conversely, assume that $EG(M) \in \mathbf{J}$. It suffices to show that every regular element of $D(M)$ is idempotent. Thus let $t$ be a regular element of $D(M)$ and let $x$ be an inverse of $t$ in $D(M)$. Then there exists a derivation

$$t_0 = \xi \to t_1 \to \cdots \to t_n = t$$

where each $t_i \in (M \cup \xi)^*$ for $1 \leq i \leq n$.

Denote by $t_i\pi$ the element of $M$ obtained by deleting every occurrence of $\xi$ in $t_i$. We claim that for $0 \leq i \leq n$,

$$x(t_i\pi)x = x$$

We prove the claim by induction on $n - i$. For $i = n$, $t_n = t$ and $x(t\pi)x = x$ by hypothesis. If $t_{i+1}$ is derived from $t_i$ by applying $\xi \to \xi\xi$ or $\xi \to 1$ then $t_{i+1}\pi = t_i\pi$ and the induction is trivial. Thus assume that $t_i = u\xi v$ and $t_{i+1} = us\xi\bar{s}v$ (the case $t_{i+1} = u\bar{s}\xi sv$ is dual). Then by induction

(1) $x(u\pi)s\bar{s}(v\pi)x = x$

Set $e = (v\pi)x(u\pi)s\bar{s}$ and $f = s\bar{s}e$. Then $e$ and $f$ are two $\mathcal{L}$-equivalent idempotents of $EG(M)$ and thus $e = f$. It follows $e(v\pi)x = f(v\pi)x$, that is, by using (1)

(2) $(v\pi)x = s\bar{s}(v\pi)x$

Now if we report (2) in (1) we obtain

(3) $x(u\pi)(v\pi)x = x$ i.e. $x(t_i\pi)x = x$

and this proves the claim.

In particular for $i = 0$, we have $t_0 = 1$ and thus $x^2 = x$. But $xtx = x$ and therefore by Proposition 4.1, $xt = tx = x$ and hence $t$ is idempotent. Therefore every regular element of $D(M)$ is idempotent and this concludes the proof. $\square$

Let us summarize the main results of this paper

**Theorem 4.7** $\mathbf{G} = \mathbf{PG} = \mathbf{J} * \mathbf{G} \subset \mathbf{J}^{-1}\mathbf{G} = \mathbf{EJ} = \mathbf{BG}$

These results should be compared with our results on the variety **Inv** generated by inverse monoids [4]. If $\mathbf{J}_1$ denotes the variety of idempotent and commutative monoids, then

$$\diamondsuit_2(\mathbf{G}) = \mathbf{Inv} = \mathbf{J}_1 * \mathbf{G} = \mathbf{J}_1^{-1}\mathbf{G} \subset \mathbf{EJ}_1$$

and we conjecture that the inclusion $\mathbf{J}_1^{-1}\mathbf{G} \subset \mathbf{EJ}_1$ is in fact an equality. Thus it is tempting to conjecture that $\mathbf{J} * \mathbf{G} = \mathbf{J}^{-1}\mathbf{G}$. However, for some reasons that cannot be discussed in this paper, our believe is that $\mathbf{J} * \mathbf{G} \subsetneq \mathbf{J}^{-1}\mathbf{G}$.

# References


[1] S. Eilenberg, *Automata, Languages and Machines*, Vol. A, Academic Press, New York, 1974. Vol. B, Academic Press, New York, 1976.

[2] J. Karnofsky and J. Rhodes, Decidability of complexity one-half for finite semigroups, *Semigroup Forum* **24**, (1982), 55-66.

[3] G. Lallement, *Semigroups and combinatorial applications*, Wiley, New York, 1979.

[4] S. W. Margolis and J.-E. Pin, Languages of inverse semigroups, in *11th ICALP*, Berlin, 1984, 337–346, *Lecture Notes in Computer Science* **172**, Springer Verlag.

[5] J.-E. Pin, Hiérarchies de concaténation, *RAIRO Informatique Théorique* **18** (1984), 23–46.

[6] J.-E. Pin, *Variétés de langages formels*, 160 p., Masson, Paris (1984).

[7] J.-E. Pin, Topologies for the free monoid, submitted.

[8] Ch. Reutenauer, Sur les variétés de langages et de monoïdes, *Lecture Notes in Computer Science* **67**, Springer, Berlin, (1979) 260–265.

[9] J. Rhodes and B. Tilson, Local complexity of finite semigroups, in *Algebra, Topology and Category Theory*, A Collection of Papers in Honor of Samuel Eilenberg, Academic Press, (1976) 149–168.

[10] H. Straubing, Varieties of recognizable sets whose syntactic monoids contain solvable groups, Ph. D., University of California, Berkeley, (1978).

[11] H. Straubing, Aperiodic homomorphisms and the concatenation product of recognizable sets, *J. Pure Appl. Algebra* **15**, (1979), 319–327.

[12] H. Straubing, Recognizable sets and power sets of finite semigroups, *Semigroup Forum*, **18**, (1979), 331–340.

[13] H. Straubing, A generalization of the Schützenberger product of finite monoids, *Theor. Comput. Sci.* **13**, (1981), 137–150.

[14] H. Straubing, Finite semigroups varieties of the form $\mathbf{V} * \mathbf{D}$, *J. Pure Appl. Algebra* **36**, (1985), 53–94.